第 54 卷 第 2 期　北京大学学报(哲学社会科学版)　Vol. 54, No. 2
2017 年 3 月　Journal of Peking University(Philosophy and Social Sciences)　Mar. 2017

从语言关联看生成词库的认知取向

宋 作 艳

(北京师范大学 文学院,北京 100875)

摘　要: 生成词库理论是自然语言意义的计算与认知模型,无论是词汇表征还是语义生成机制都融入了认知视角。相比之下,生成词库更加注重认知的语言关联,通过物性结构将一些重要的百科知识引入词义描写,建立了百科知识到语言知识的接口,丰富了名词的词汇表征。这些百科知识在语义生成(组合)中起重要作用。物性结构在范畴化、隐喻、转喻和概念整合等认知语言学研究领域得到广泛应用,可以进一步揭示百科知识的语言关联,同时说明生成词库与认知语言学有相通之处。百科知识的语言关联在不同语言中有普遍性,也有相对性,反映了人类语言和认知的共性与差异。

关键词: 生成词库理论;认知语言学;物性结构;百科知识;语言关联

中图分类号:H 0　**文献标识码:**A　**文章编号:**1000-5919(2017)02-0151-10

生成词库理论(Generative Lexicon Theory)是由普斯特若夫斯基(Pustejovsky)于 20 世纪 90 年代初创立的,经过不断完善,已成为语义学研究的范式之一,广泛应用到各种语言的研究中。

生成词库理论是面向语义计算的,关注词义的形式化分析,被称为"当代语义学中最精细的形式化分解模型"①。实际上,它是自然语言意义的计算与认知模型,认知性也是它的一个突出特点。与认知语言学一样,生成词库理论主张意义本质上是概念性的,但并不认为所有百科知识都属于词义,而是通过物性结构(qualia structure)在词义描写中引入了部分具有语言关联的百科知识(encyclopedic knowledge),并展示了这些知识在语言中的表现,从而揭示了认知的语言关联。本文旨在通过大量语言事实来分析生成词库理论的这一认知取向特点,首先简单介绍这一理论的核心思想及其认知表现,之后重点讨论物性结构,最后通过物性结构在认知语言学研究中的应用来进一步证明百科知识的语言关联。

一、生成词库理论的核心思想及其认知表现

生成词库理论首次把广义的生成方法从语法扩展到词库,其基本观点是,词的义项是相对稳定的,但在组合中是动态的、有弹性的,可以通过语义生成机制生成很多具体的意义解释。因此,要限制词库中义项的数量,通过丰富词项的词汇表征(lexical representation)和语义生成机制(generative mechanisms in semantics),捕捉词在上下文中潜在的无限意思②。词项的词汇表征包括四个层面:论元结构(argument structure)、事件结构(event structure)、物性结构(qualia structure)和词汇类型结构(lexical typing structure)。其中物性结构包括构成角色(constitutive role)、形式角色(formal role)、功用角色(telic role)和施成角色(agentive role),分别说明词项所指的事物由什么构成、属于什么类、有什么用途或功能以及怎样产生的。语义生成机制最初有类型强迫(type coercion)、选择性约束(selective binding)和协同组合(co-composition)三种③。后来,基于论元选择,语义生成机制改进成纯粹类型选择(pure

收稿日期:2016-11-20
作者简介:宋作艳,女,山东淄博人,北京师范大学文学院副教授。
基金项目:教育部人文社会科学研究青年基金项目"汉语名词语义—语法双向互动关系研究"(项目批准号:15YJC740069);中央高校基本科研业务费专项资金;国家社科基金重大项目"基于中国语言及方言的语言接触类型和演化建模研究"(项目批准号:14ZBD102)。

① Geeraerts, D., *Theories of Lexical Semantics.* Oxford: Oxford University Press, 2010, p. 147.

② Pustejovsky, J., *The Generative Lexicon*, Cambridge, MA: MIT Press, 1995, p. 105.

③ Ibid., pp. 105-140.

selection)、类型调节(type accommodation)和类型强迫①。语义生成机制通过选用词项(尤其是名词)词汇表征中的信息来实现语义生成。

以 finish(完成)为例,finish the book 的默认意思是 finish reading/writing the book(读完/写完这本书),finish the beer 意思是 finish drinking the beer(喝完这啤酒),finish the meal 意思是 finish eating the meal(吃完这饭)。其中的 finish 词义一样,只是在上下文中会生成具体的意思,隐含的动词都可以在名词的物性结构中找到:read 和 write 分别是 book 的功用角色和施成角色;drink 是 beer 的功用角色;eat 是 meal 的功用角色。再比如,“好书”是“好看的书”,“好饭”是“好吃的饭”。可见,词义在上下文中是可协商的(negotiable),但由语义生成规则控制,而不是语用推理。通过丰富的词汇表征和有限的语义生成机制,生成词库理论把意义的变化放到组合层面去解决,“将意义负担均匀地扩散到了组合中的其他要素上”②,尤其是名词上,从而减轻了动词的语义负担,不仅避免了过多的多义处理,而且把词义与短语、句子的意义联系在了一起。需要注意的是,生成词库理论与生成语法中的“生成”含义不同:前者生成的是合法的语义,关注语言的理解与解释;后者生成的是合法的语言形式,关注语言的生成。

与义项列举相比,语义生成模型更符合人脑对语言运作机制的认知。人脑不同于电脑,受记忆力的限制,心理词库(mental lexicon)中不可能储存太多词义。不可能像传统语义学中的义项列举法(sense enumeration)那样为每个词列举多个义项,如为 finish 列举 finish reading、finish writing、finish eating 等义项。这种静态的词义描写方法是机械的、被动的,不仅会造成词库的极度膨胀,而且无法穷尽词在上下文中的所有变化③。那么,我们如何在无限的语境中有效利用有限的词汇呢? 生成词库理论认为词义组合成短语义、句义的过程是动态的、积极的,是协同组合,词(尤其是做论元的名词)会调整自己的意义来适应组合的要求。这样就可以控制义项的数量,减轻词库的负担,在具体的语境中通过语义生成机制和开发利用词义中的信息来生成具体的语义解释。语义生成从词的角度看是词义在上下文中进行自我调节,生成具体的语境义;从组合(短语、句子)的角度看是词义通过自我调节协助生成组合的具体意思,即意义具体化、清晰化的过程。从认知角度看,词义生成就是凸显、激活词义的某一部分。这样的语义生成模式更突出认知的积极作用,通过语义信息和语境信息之间的接口,展现了词汇和认知之间的互动关系。

不过,生成词库理论的认知性主要源于它对意义本质的认识。生成词库理论认为,人类的认知能力反映在语言中,尤其反映在心理词库中④;词义要能反映认知系统中更深层的概念结构,也就是说,自然语言语义学应该是非语言学的概念组织原则的映像(image)⑤。这一观点遵循了认知语言学的基本假设:意义是认知模式中的概念化,存在于人的大脑中,是从语言的表达式向某些心理活动的映射⑥。这意味着生成词库理论不只是语言学理论,还是关于人类认知结构的理论。与意义本质问题密切相关的是关于语言知识与百科知识的区分,意义是概念性的主张直接影响了生成词库的词汇表征,使之通过物性结构把部分百科知识纳入到了词义中,并探讨了这些百科知识在语义生成中的作用,揭示了认知的语言关联。

---

① 参见 Pustejovsky, J, “Type theory and lexical decomposition”, *Journal of Cognitive Science*, 2006, 6:39 – 76; Asher, N. and J. Pustejovsky, “A type composition logic for generative lexicon”, *Journal of Cognitive Science*, 2006, 7.1:1 – 38.

② Pustejovsky, J, *The Generative Lexicon*, p. 113.

③ Ibid., pp. 59 – 60.

④ Pustejovsky, J, “Type construction and the logic of concepts”, in P. Bouillon and F. Busa (eds.), *The Language of Word Meaning*, Cambridge: Cambridge University Press, 2001, pp. 91 – 123.

⑤ 参见 Pustejovsky, J, “The Generative Lexicon”, *Computational Linguistics*, 1991, 17.4:409 – 441; Pustejovsky, J, *The Generative Lexicon*, p. 6.

⑥ Gärdenfors, P, Some tenets of cognitive semantics, in J. Allwood and P. Gärdenfors (eds.), *Cognitive Semantics – Meaning and Cognition*, 19 – 36, Amsterdam: John Benjamins, 1999.

二、物性结构:认知的语言关联——从百科知识到语言知识

生成词库理论继承了分解语义学(decompositional semantics)的理念,认为词义可以分解成一系列语义成分或语义特征,这些特征在词义的组合中会起作用。这种方法本质上是认知性的,反映了我们对词义的理解。但与以往的词汇分解理论不同,生成词库在词义分析中引入了百科知识。

结构主义语义学中的语义成分分析法严格区分语言知识(linguistic knowledge)和百科知识。在结构主义看来,词义并不直接对应客观世界的任何事物,也不直接对应概念,而是取决于词在语言系统中的位置。心理词库包含词的定义,使之可以映射到它们所表征的概念上,而概念表征与词的指称相关的更丰富的百科知识①。语义成分分析要求必须符合必要充分条件(necessary and sufficient condition),即语义成分分析的结果与义位的意义相等。比如,把"女人"分解成[+人+女性+成年]。至于女人的其他特征,比如"温柔、贤惠"等都不会归为语义成分,因为这些信息都是与词的指称相关的百科知识②,即人类经验世界中关于事物的知识。

生成语义学继承了词汇分解法,并且排斥百科知识。而之后的自然语义元语言理论、概念语义学和生成词库理论打破了意义的词汇表征与概念表征之间的界限,开辟了面向更广泛认知背景的意义分解的道路,关注词库与广义认知的互动关系,寻找词义分解的认知基础,考察语言语义学与非语言信息之间的接口③。是否在词义中引入百科知识,取决于对意义本质的认识。这三种理论都认为,意义本质上是概念性的,而这恰恰是认知语言学的重要信条之一。认知语言学认为,意义本质上是一种认知现象,意义就是概念化,词语的意义等同于说话者或听话者大脑中被激活的概念,与人类的一般认知能力和方式具有密切的关系,要从一般认知结构来描写语义④。因此,认知语言学持"百科全书式"的语义观,认为语言知识与百科知识没有明确的界限,词义具有百科知识的特点,和概念有关的任何知识都是词义的一部分⑤。因为关于事物的任何层面的知识,原则上都可能在确定表达这一事物的词语的语言行为时起到一定作用。

语言知识与百科知识的区分问题与词汇语义分析的心理现实性与充分性、简约主义与繁复主义纠缠在一起。排斥百科知识符合简约主义,但会影响意义表征的心理现实性和充分性。词义表征不充分,词在组合中的有些表现就很难解释,比如对"很女人"的理解必须引入女人温柔、贤惠等知识。吸纳百科知识,就势必要走繁复主义的路子,走向认知。生成词库理论要解释词在组合中的生成性就要丰富词项的词汇表征,丰富词汇表征就要面临语言知识与百科知识的界限问题。普斯特若夫斯基(Pustejovsky)一方面承认语言知识与百科知识之间是连续的,另一方面坚持要区分二者,认为只有部分百科知识属于名词词义,需要在词义表征中描写,会在意义组合中起系统性作用⑥。生成词库理论既不像语义成分分析法和生成语义学那样保守,完全排斥百科知识,也不像认知语言学那么激进,把一切百科知识都纳入词义,而是通过物性结构等将一部分会影响语义生成的百科知识引入到词义描写中。

---

① Murphy, M. L., *Lexical Meaning*, Cambridge: Cambridge University Press, 2010, p.45.

② 又称为"概念知识"(conceptual knowledge)、"世界知识"(world knowledge)和"常识知识"(commonsense knowledge)。

③ 参见 Geeraerts, D, *Theories of Lexical Semantics*, p.125; Murphy, M.L., *Lexical Meaning*, p.45.

④ 参见 Langacker, R.W, *Foundations of Cognitive grammar*, vol.1, Stanford: Stanford University Press, 1987; Langacker, R.W., *Grammar and Conceptualization*, Berlin: Mouton de Gruyter, 1999; Talmy, L, *Toward a Cognitive Semantics* (vol.1): *Concept Structuring Systems*, Cambridge: MIT Press, 2000. 与生成语法不同,认知语言学认为语言能力没有独立的认知模块,属于一般认知能力。

⑤ 参见 Haiman, John., Dictionaries and encyclopedias, *Lingua*, 1980, 50:329-357; Langacker, R.W, *Foundations of Cognitive grammar*, vol.1, Stanford: Stanford University Press, 1987; Langacker, R.W., *Grammar and Conceptualization*, Berlin: Mouton de Gruyter, 1999; Croft, W., The role of domains in the interpretation of metaphors and metonymies, *Cognitive Linguistics*, 1993, 4.4: 335-370.

⑥ Pustejovsky, J., *The Generative Lexicon*, p.4, p.43, pp.232-236.

物性结构源于亚里士多德的“四因说”(Aristotle's four causes),包括四种物性角色(qualia roles),提供四个不同方面的信息:

1)构成角色:描写事物与其组成部分之间的关系。包括材料、重量、成分和组成部分。

2)形式角色:描写事物在更大的认知域内区别于其他事物的属性。包括方位、大小、形状和维度等。

3)功用角色:描写事物的用途和功能。

4)施成角色:描写事物怎样形成或产生的,如创造、因果关系。

以“书”为例,它的构成角色包括“封面、封底、信息”等,形式角色说明“书”的上位是物体和信息,还包括形状、尺寸、颜色等特征,功用角色是“读”,施成角色是“写”。此外,规约化属性(conventionalized attribute)也被纳入到物性结构中①,指人类经验中可以感知到的事物的典型特征,比如“叫”是“狗”的规约化属性,“流动”是“水”的规约化属性。不是所有名词都含四种角色,比如指自然物的名词通常没有施成角色。物性结构描写了名词中典型的谓词和关系,把词与概念网络联系起来,是概念逻辑的组织原则。在认知语言学看来,物性结构就是观察视角(Ways-of-Seeing)②:不同物性角色分别对应部分—整体、类别、功能和生命史四种观察视角。兽医、动物学家和骑师会分别从部分—整体、类别和功能的角度来看马,作者会从生命史的角度来看书。理论上,任何词都可以有物性结构,但实际上,物性结构更适合用来分析名词语义。物性结构之于名词,就像论元结构之于动词。物性结构决定了名词的语义,就像论元结构决定了动词的语义③。物性结构的引入丰富了名词的意义表征,大大推进了名词语义的研究,直接导致了从动词视角到名词视角的转向。

在词义分析中引入百科知识,生成词库理论并非最早,它的进步之处在于将这些知识结构化(structurize)了,组织成经验的概念域(experiential domain),而且只引入部分百科知识。这些知识是系统的、受约束的,在语言的生成和理解中起各种作用,不同于语用学中不受约束的百科知识④。换言之,这是词汇中固化的知识,比如酒通常是用来喝的,在一定情况下偶尔会用它来消毒,但“消毒”不是“酒”的词义的一部分,“喝”才是,可用于“好酒”“酒后”等组合的语义理解。简而言之,这些知识都有语言关联(linguistic relevance),会影响语言表达的及其理解,是语言中能够区分的知识或语言形式和内容能够对应起来的知识,可以成为语言规则的条件⑤。只有符合语言关联原则的百科知识才属于语言知识,才属于词义,需要在物性结构中描写。生成词库通过物性结构为百科知识和语言知识提供了接口,揭示了百科知识的语言关联,也就揭示了认知的语言关联。事实证明,不同物性角色都在语言生成或理解中起作用。比如,形式角色中形状义凸显的名词常会发生隐喻引申,功用义凸显的名词(比如工具名词)更倾向于用作动词(详见下文)。物性结构是词库生成性特征的核心,不仅结构化了我们的词汇知识,而且暗示了词在上下文中的解释⑥。如,book 的功用角色 read 和施成角色 write 可以为 finish the book 提供默认解释 finish reading/writing the book;“书”的功用角色“写”可以为“赶论文”提供默认解释“赶写论文”;“钢琴”的功用角色“弹”可以为“学习钢琴”提供默认解释“学习弹钢琴”。这些隐含的信息以往都是放在语用推理中处理,生成词库将它们纳入词义,在物性结构中描写,供语义生成时调用。普斯特若夫斯基(Pustejovsky)更关注物性角色在语义生成中的作用,尤其是多义和语义模糊现象⑦。

---

① Pustejovsky, J. and E. Jezek, “Semantic coercion in language: beyond distributional analysis”, *Italian Journal of Linguistics*, 2008,20.1: 181 – 214.

② Croft W. and D. A. Cruse, *Cognitive Linguistics*, Cambridge: Cambridge University Press, 2004, pp.137 – 138.

③ 参见 Pustejovsky, J., “The Generative Lexicon”, *Computational Linguistics*, 1991, 17.4:409 – 441; Pustejovsky, J. and B. Boguraev., “Lexical knowledge representation and natural language processing”, *Artificial Intelligence*, 1993, 63: 193 – 223.

④ Jackendoff, R, *The Architecture of the Language Faculty*, Cambridge, MA: MIT Press, 1997, pp.61 – 64.

⑤ 陈保亚:《从百科知识到语言知识——认知的语言关联及其相对性》,《贵州民族大学学报》2015 年第 2 期。

⑥ Pustejovsky, J., *The Generative Lexicon*, p.87.

⑦ 不同物性角色的语言关联详见 Pustejovsky(1995)。

三、物性结构与认知语言学研究

生成词库通过物性结构引入百科知识,丰富了词汇的意义表征,为如何分析百科知识和语言规则的关系提供了典范。设立物性结构的最初目的是为了抓住词义在上下文中的变化,实际上,物性结构中的百科知识已经远远超出预期的目标,广泛应用到范畴化、转喻、隐喻与概念整合等认知语言学相关问题的研究中。这既说明生成词库理论与认知语言学有相通之处,也进一步证明了这些百科知识有更多的语言关联,有更高的语言价值,可以转化为语言知识。

1. 范畴化

范畴是我们理解和认识世界的基础,范畴化是对世界上的事物进行分类的过程,是人类经验和想象的结果,一方面来源于感知、行为活动以及文化的交互作用,另一方面来源于隐喻、转喻以及心理意象等认知途径①。物性结构中的四种物性角色“推动了我们对世界中事物或关系的理解,进一步有助于(实际上可能决定了)我们命名事物的能力”②。换言之,物性结构中的特征反映了人类对世界的感知和认知,为范畴化提供经验基础,是人类对经验世界中的事物进行分类的重要依据,至少对指具体事物的名词而言是这样的。克拉克和克拉克(Clark and Clark)提出了事物范畴化的基础:物理特征(physical characteristics)、个体发生史(ontogeny)和潜在作用(potential role)③。物理特征大致对应形式角色和构成角色,个体发生史对应施成角色,潜在作用对应功用角色。

物性结构在范畴化中的作用可以由三方面的证据来证明,一是词典释义。我们描述新事物,或者定义事物的时候,一般都从物性结构中的特征入手。以“砖”的释义为例④:

砖(1) [名] 把黏土等做成的坯放在窑里烧制而成的建筑材料,多为长方形或方形。

“黏土”说明砖的材料,属于构成角色;“烧制”说明砖的制作方式,属于施成角色;“建筑材料”说明砖的用途,属于功用角色;“长方形或方形”描述砖的形状,属于形式角色。相关的范畴常常以这些特征互相联系与区别,比如“瓦”也是建筑材料,也是黏土烧制而成的,与“砖”的不同主要在具体功用和形状上:瓦通常用来铺屋顶,多是拱形的、平的或半个圆筒形的等。

二是定中复合名词或短语中的修饰关系。修饰性成分通常描述事物的颜色、形状、功用、产生方式、典型特征等,与物性结构的角度一致。例如,“黑板”与“石板”的修饰成分分别属于形式角色(颜色)和构成角色(材质);“纸杯”与“茶杯”的修饰成分分别属于构成角色(材料)和功用角色(喝茶);“刀背、刀把、刀刃”的整体—部分构成关系不同;“菜刀、砍刀、刺刀、铅笔刀”的功用不同;“雕像、绣像、塑像”的制作方式不同;“飞鸟”“流水”与规约化属性相关。形容词也能体现四种修饰关系,例如,“白饭”指“不加菜、糖等做成并且不就菜吃的米饭”,与“饭”的构成角色相关;“白布”与“布”的形式角色(颜色)相关;“庸医”指“医术低劣的医生”,涉及“医”的功用角色(治病);“旱稻”的意思是“旱地里种植的水稻”,涉及“稻”的施成角色(种)⑤。意大利语中的定中名词短语用介词 da 标记功用修饰关系,di 标记施成修饰关系,a 和 al 标记构成修饰关系⑥:

(1) bicchiere　da　vino
杯子　介词 - da　酒　(酒杯)

---

① 束定芳编著:《认知语言学》,上海:上海外语教育出版社 2008 年版,第 67 页。

② Pustejovsky, J., *The Generative Lexicon*, p. 85.

③ Clark, E. V. and H. H. Clark, “When nouns surface as verbs”, *Language*, 1979, 55.4: 767 - 811.

④ 文中释义均引自《现代汉语词典》(第 6 版)。

⑤ 张念歆、宋作艳:《汉语形名复合词的语义建构:基于物性结构与概念整合理论》,《中文信息学报》2015 年第 6 期。

⑥ 引自 Johnston, M. and F. Busa, Qualia structure and the compositional interpretation of compounds, in E. Viegas (ed.), *Breadth and Depth of Semantics Lexicons*, Dordrecht: Kluwer, 1999, pp. 167 - 187.

(2)foro di pallottola
孔 介词 – di 子弹 (弹孔)
(3)porta a vetri
门 介词 – a 玻璃 (玻璃门)
(4)seni al silicone
乳房 介词 – al 硅胶 (硅胶假胸)

这说明,百科知识(物性特征)在不同语言中的语言关联方式不同,关联强度也不同,反映了人类的语言差异和认知差异。在意大利语中,三种知识的区分有语言标记,语言关联性更强,更有语言价值。

三是分类词(classifier)。对名词和名词性成分进行分类的手段统称为分类词,分类词是观察人类如何建构世界表征、并将之编码为语言的重要窗口。无生名词的分类词主要根据物理性质和功能来区分①,对应名词形式角色、构成角色和功用角色中的特征。其中,物理性质包括外形、维度、方向、大小、状态、材质等。汉语中的数分类词(即量词)主要是根据名词的外形来区分的,如用于条形物的“根、条、道、支”,用于平面的“面、片、张、幅”,用于圆形的“颗、粒、丸”(后两者只能用于“小”的东西),用于花状事物的“朵”,用于盘状事物的“盘”。有些语言中有表示木头或金属物体的分类词。澳大利亚的 Yidiny 语有一整套功能分类词,分别是 minya“可食的肉类食物”、mayi“可食的非肉类食物”、bana“可喝的液体”、bulmba“可居住的地方”、gugu“故意制造的声音”、wirra“可移动的物体”②。不同语言的分类词选用的特征重叠性很强,同时又不完全一致,说明这些特征具有普遍的语言价值,其语言价值又因语言不同而异,说明百科知识的语言关联有普遍性也有相对性。

范畴化往往会融入人的目的性和主观意图,这一点集中体现在含功用角色的范畴中。在四种角色中,功用角色处于核心地位,构成、形式和施成都是为功用服务的。砖用什么做,做成什么样子、怎样制作都是为了更好地实现其功能。功用义对名词词义和构词有系统性影响,具有重要的语言价值和语言学价值③。其中之一就是影响名词的范畴化,无论是上位范畴还是下位范畴,都常以功用特征作为范畴化的基础。人造物都是为了一定目的而制造的,比如“书”是为了“读”而写的,因此指人造物的名词都有固有功用义。有些概念可以指自然物,但其范畴化完全基于功用,例如“玩具”(专供玩儿的东西)、“食物”(可以充饥的东西),再如“奶牛”(主要为产奶而饲养的牛)、“口条”(用作食品的猪舌或牛舌)。功用修饰模式是汉语中很能产的复合词构造方式,而且广泛存在于各种语言的复合词中,如汉语中的“烤箱、浴室、书桌、纸篓、打印机、豆浆机、粉碎机、复印机”,英语中的 drinking water(饮用水),日语中的“本箱”(书箱),德语中的 briefmarke(letter-mark 邮票)。土耳其语中功用的语言关联性更强,功用修饰要么在中心名词上加后缀 – i 或其语素变体,要么在定语名词后加后缀-lik 或其语素变体④。职业名词和工具名词的命名常常是从功用角度,如“管家、导演、编辑、刹车、刷子、锯”。功用是一种社会属性,与具体的社会文化有关,不同的语言会有差异。比如,澳大利亚人把狗看成亲密的朋友,不会把“狗肉”范畴化为“食物”。再比如,英语不区分指自然部件的“舌头”和做食物的“口条”,都用 tongue。却区分自然部件 flesh 和供食用的肉 meat,汉语都用“肉”。这同样说明百科知识的语言关联具有相对性,会因语言而异,背后的根源是认知的相对性和文化的差异。

在生成词库中,“石头、草、男人”等是自然类(natural type),“砖、纸、医生”等是人造类(artifactual

---

① Aikhenvald, A. Y, *Classifiers: A Typology of Noun Categorization Device*, New York: Oxford University Press, 2000, pp. 271 – 305.

② 引自 Dixon, *Where Have All the Adjectives Gone? And Other Essays in Semantics and Syntax*, Berlin: Mounton, 1982, pp. 191 – 205.

③ 详见宋作艳:《功用义对名词词义与构词的影响——兼论功用义的语言价值与语言学价值》,《中国语文》2016 年第 1 期。

④ 引自 Bassac, C. and P. Bouillon, The telic relationship in compounds, in J. Pustejovsk *et al.* (eds) *Advances in Generative Lexicon Theory*. Springer, 2013, pp. 109 – 126.

type)。两类之间最大的区别是后者融入了人的意图和目的,含功用角色。区分自然类的是偶分的自然属性,类之间彼此排斥;区分人造类的功用属性则是多分的,会造成语义类的多重继承。例如,“沙发床”既是沙发也是床,而“苹果梨”是形状像苹果的梨,不是苹果。从自然属性看,“狗”的上位词只能是“动物”,但从功用属性看,它的上位可以是“家畜”“宠物”等。“刀”可能是厨具、兵器,也可能是凶器,跟前两者的关联显然更大。

转喻范畴化通常与构成角色、功用角色有关,比如“须眉”转指“男子”涉及构成角色(部分转指整体);“鱼”转指“鱼肉”与鱼肉是常见食物有关。隐喻范畴化也以物性角色为基础,“纸花、山脚、笔帽、海米”中的隐喻分别基于形状、构成、功用和制作方式的相似性(详见下文)。

2. 转喻

转喻是一般的认知方式,基于事物之间的邻近性(contiguity),用一个事物来指代另一个邻近的事物。邻近性源于经验世界中事物之间的相关性①,而相关性体现在物性结构中。生成词库理论中的类型强迫机制解决的就是转喻问题,为转喻提供了形式化的解释:通过强迫论元转换类型来满足谓词的要求,类型的转换借助物性结构中的信息。类型强迫处理的一个重要现象就是逻辑转喻(logical metonymy),即名词短语被用在了动词短语的位置,事物转指与之相关的事件。例如:

(5)John began the book.

begin 要求带事件类论元,而 the book 不合要求,begin 就强迫类 the book 转换成事件 to write/to read the book,实现转喻重建;类型转换是借助 book 的施成角色 write 和功用角色 read 完成的②。可见,物性结构连接事件结构与论元结构,即动词的动态语义和静态语义,为抽象的框架提供内容。

名词动用也属于逻辑转喻,可以用生成词库理论中的类型强迫来解释③。名词动用的本质是以事件中的参与者来转喻事件,至于是哪些事件,以往的研究多是根据名词在动词的论元结构中承担的语义角色来分类,通常分为 7 – 10 类,包括工具、材料、处所、地点、覆盖物、填充物、结果、对象、施事等,这是以动词为中心的研究视角。如果反过来,从名词出发,看动词在名词的物性结构中充当什么物性角色,可以把这些事件合并成三类,分别与名词的功用角色、施成角色和规约化属性有关。例如,“百度一下,你就知道”中的“百度”转指“用百度搜索”,“搜索”是“百度”的功用角色。显然,这一视角的解释概括性更强,而且能回答“什么样的名词倾向于活用为动词”:具有明确、单一功用角色或施成角色的名词更易用作动词。即人造类名词更易用作动词,比如工具名词。而且,这种解释更准确,比如,“把白菜窖上”并非只是“把白菜放到窖里”(“窖”是“放”的处所),而是“把白菜储藏到窖里”(“储藏”是“窖”的功用)。关注的是“窖”的功用义,而不是处所义。

概念转喻在生成词库理论中也是通过类型强迫和物性结构信息来实现的,比如“他开了一辆奔驰”是下位类“奔驰”转指上位类“车”的子类强迫(subtype coercion),与形式角色中的上下位关系有关;“我喜欢读莫言”是用“生产者”(莫言)转指产品(莫言写的书),与“书”的施成角色有关。生成词库减少义项数量的方法之一是把一个词的不同义项合并到一个元词条(meta-entry)中,称之为词汇概念聚合(lexical conceptual paradigm)④。这些义项之间彼此有联系,分属不同的语义类,一起构成合成类(complex type)。合成类反映了概念化与词义之间的关系以及人类观察事物的方式,与认知语言学中的多面义(facet)大致相当⑤。“书”就是一个合成类[物体 · 信息],两个子类之间是承载关系。相比之

---

① Lakoff, G. and M. Johnson, *Metaphors We Live By*, Chicago: University of Chicago Press, 1980, pp. 35 – 40.

② Pustejovsky, J. and P. Bouillon, “Aspectual coercion and logical polysemy”, *Journal of Semantics*, 1995, 12. 2: 133 – 162.

③ 详见宋作艳:《逻辑转喻、事件强迫与名词动用》,《语言科学》2013 年第 2 期;宋作艳:《生成词库理论与汉语事件强迫现象研究》,北京:北京大学出版社 2015 年版。

④ Pustejovsky, *The Generative Lexicon*, p. 91.

⑤ Croft W. and D. A. Cruse, *Cognitive Linguistics*, Cambridge: Cambridge University Press, 2004, pp. 116 – 126.

下,“小说”就只有信息义。不同的语境会选择合成类的不同子类或整体来生成具体意思,例如,“尽信书则不如无书”选择“书”的信息义,“焚书坑儒”选择物体义,“有趣的立体书”两个意思都有。合成类处理的是一些固化的转喻,比如“报纸”是合成类[机构·物体·信息],“机构”是“报纸”的出版单位,涉及“报纸”的施成角色。整体—部分转喻是概念转喻中常见的类型,不仅与构成角色有关,还与功用角色有很大关系①。如“香蕉”既指一种植物,也指这种植物的果实,属于合成类[植物·果实],在词典中分列两个义项。这类多义词大量存在,而且成系统性。有的转指种子,如“玉米”“小麦”;有的转指茎,如“甘蔗”;有的转指花,如“菊花”“玫瑰”;有的转指木料,如“紫檀”。转指哪个部件取决于哪个部件功用更突显。不同语言在合成类上有相同之处也有差异,反映了人类认知和概念化的共性和个性。比如,“苹果”和 apple 都是合成类[植物·果实];rabbit 既指兔子,也指兔毛,汉语中的“兔子”却不能指兔毛。

3. 隐喻

生成词库理论重点放在与转喻相关的多义现象上,而没有关注对隐喻的研究②。实际上,物性结构也可以用于隐喻分析③。隐喻是人类用来组织概念系统不可缺少的认知工具,是通过一种事物来理解另一种事物的手段。隐喻的运作方式是不同概念域(domain)之间的映射,隐喻产生的基本条件是相似性,物性结构可以揭示相似涉及的特征范围。

统计发现,双音节隐喻名名④复合词涉及的相似性特征与物性角色有对应关系,涉及最多的是形式角色,其次是构成角色和功用角色,涉及施成角色的极少,如“海米”用“米”的加工过程隐喻“虾仁”的制作工序⑤。

形式角色因其空间性,最易被感知,因而最易作为理解其他陌生或抽象概念的基点。其中形状相似最常见,基于外形相似的后隐喻定中复合词是汉语中非常能产的一种构词模式,如“雪花、浪花、水花、烟花、火球、雪球、棉球”中的“花”“球”分别指形状像花的东西和形状像球的东西⑥。名词不仅可以通过隐喻来指称事物,还可以通过隐喻来描述事物的特征,这一功能源于名词丰富的物性特征。例如,“茶砖”是用“砖”指称形状像砖的东西,“砖茶”是用“砖”的形状来描述“茶”的形状。另如“球果”“雪柳”和“鸡冠花”都涉及隐喻性描述。

与构成角色相关的隐喻主要涉及部分与整体关系的相似。例如,“碑首”基于“碑的顶部—碑”与“首—人体”两种部分—整体之间关系的相似性;“本末”是用树根和树梢来比喻事情的头尾,涉及部分与部分之间关系的相似性。这类隐喻常涉及器官类名词,如“耳房、脚注、眉批、山脚、页眉”。

与功用角色相关的隐喻主要涉及人造类名词和器官名词,依据的是功能的相似性,常常同时涉及形状的相似性。如“桌子腿”中的“腿”指“器物下部像腿一样起支撑作用的部分”;“笔帽”中的“帽”指“罩或套在器物上头,作用或形状像帽子的东西”。

词中的隐喻涉及隐喻范畴化,都已固化。句子中有些临时性隐喻的理解也需要借助物性角色。在下面这个句子中,“栋梁”的比喻义已经固化了,用支撑房屋的大梁来“比喻担负国家重任的人”,基于功用相似性;“砖瓦”和“沙子”的比喻义是临时性的,也是就功用而言的。

(6)为人倘不能成为支撑大厦的栋梁,则甘心做平凡的砖瓦和沙石。

① 详见宋作艳:《功用义对名词词义与构词的影响——兼论功用义的语言价值与语言学价值》,《中国语文》2016 年第 1 期。

② Pustejovsky(1995:267)已经提到这一点。

③ 参见黄洁:《汉英隐转喻名名复合词语义的认知研究》,《外语教学》2008 年第 4 期;宋作艳、赵青青、亢世勇:《汉语复合名词语义信息标注词库:基于生成词库理论》,《中文信息学报》2015 年第 3 期;李强:《基于物性结构和概念整合的名词隐喻现象分析》,《语言教学与研究》2014 年第 6 期。

④ 两个名语素都指具体事物。

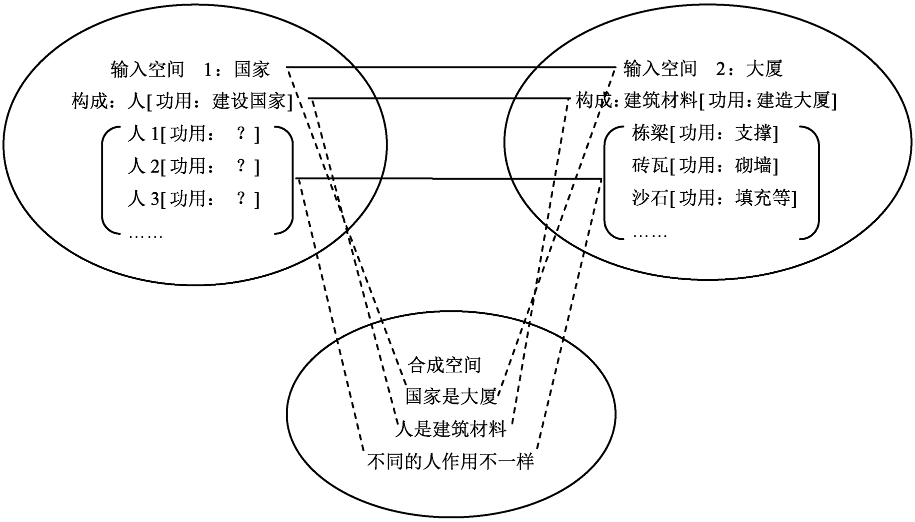
⑤ 详见赵青青:《现代汉语名名复合词研究——基于生成词库理论》,北京师范大学硕士论文,2014。

⑥ 详见宋作艳:《定中复合名词中的构式强迫》,《世界汉语教学》2014 年第 4 期。

4. 概念整合

概念整合(Conceptual Blending)是创建心理空间(mental space)之间连接网络的过程,有选择地将两个输入空间(input space)中的内容投射到合成空间(blend space)中去,实现意义的建构①。具体整合哪些内容需要具体词汇来提供,名词的物性结构可以为输入空间提供知识参考框架,从而使概念整合理论更具有操作性。物性结构与概念整合结构互为补充,一个提供分解的具体信息,一个提供抽象的整合框架,二者形成有机的杂合模式,特别适合解释组合和隐喻的意义建构②。

选择不同的物性角色或物性角色值决定了输入空间到整合空间的不同投射,从而导致对组合意义的不同解读。如"白饭"和"白布"分别选择名词的构成角色和形式角色。再比如,"红笔"通常指"笔芯"是红色的笔。在一定的语境下,也可以指"笔帽""笔壳""笔水"等是红色的笔,四者都是"笔"的构成角色。之所以默认是红笔芯,是因为笔芯与笔的功用"写"关系更密切。例(6)的隐喻也可以用物性结构和概念整合来获得识解(construal)。如下图所示,这个句子的理解需要"国家"和"大厦"两个输入空间,"国家"是由很多"人"组成的,人的共同目标是建设国家;大厦是由栋梁、砖瓦、沙石等组成,它们的共同用途都是建造大厦,但具体的用途不同,重要性不同。由两个空间的映射关系可以得到合成空间的解释:国家是大厦,人是建筑材料,每个人的作用不一样——有的人负担国家重任,起支撑作用;有的人负担一般任务,起协助作用。整个句子是建筑隐喻,涉及"大厦"的构成角色及其构成角色的功用信息。



四、结论与余论

生成词库得益于谓词逻辑以及词汇和概念结构之间关系的深入思考,通过物性结构将一些重要的百科知识引入到词义描写中,将之结构化,并通过大量语言事实展示了这些百科知识的语言关联,揭示了认知的语言关联。物性结构丰富了名词的意义表征,使语言研究从动词视角转向名词视角,并在认知语言学研究领域中得到应用,为范畴化提供经验基础,为转喻提供相关特征,为隐喻提供相似特征,为概念整合提供知识参考框架。百科知识的语言关联既有普遍性也有相对性,既与语言的类型有关,也反映了人类认知的共性与差异。相对性表现在,不同语言可能选用不同的百科知识,对相同的百科知识可能有不同的处理,相同的百科知识在不同语言中的关联方式和关联强度不尽相同。比如,Yidiny 语中有一

① 参见 Fauconnier, G. and M. Turner, "Conceptual integration networks", *Cognitive Science*, 1998, 22.2 : pp. 133 – 187; Fauconnier, G. and M. Turner, *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*, New York: Basic Books, 2002.

② 参见张辉、范瑞萍:《形名组合的意义建构:概念整合和物性结构的杂合分析模式》,《外国语》2008 年第 4 期;李强:《基于物性结构和概念整合的名词隐喻现象分析》,《语言教学与研究》2014 年第 6 期;张念歆、宋作艳:《汉语形名复合词的语义建构:基于物性结构与概念整合理论》,《中文信息学报》2015 年第 6 期。

整套功能分类词,土耳其语中有专门的功用词缀,功用在这两种语言中的语言关联性都很强,虽然具体关联方式不同。百科知识的语言关联性越强,语言价值就越大,越倾向于转化为语言知识。

本文认为,生成词库认知取向的最大特点是注重认知的语言关联,希望只把具有语言关联的百科知识通过物性结构纳入词义,但物性结构中的知识是否都具有系统的语言关联并未得到证明。物性结构来源于亚里士多德的四因说,其出发点是探讨事物存在和变化的原因,后来莫拉维斯克(Moravcsik)将其作为事物的描述模式,帮助我们了解事物①,普斯特若夫斯基(Pustejovsky)进一步将其引入词义。也就是说,物性结构中的知识从根本上说是本体性的,并非经过语言关联性验证后筛选出来的;物性结构的语言关联只是一种假设,并未得到充分证明。这也是为什么生成词库只是给出了物性结构的大致框架,至于物性结构到底要包括哪些百科知识,这些知识在语言中都有哪些表现,尚未有很好的证明。本文也只是笼统地谈物性结构的语言关联和语言价值,实际上,每种物性角色及其具体值的语言关联强度不一样,表现也不一样。相比之下,功用、形状、材料、构成等知识在语言中的作用更大,语言关联性更强,更倾向于成为语言知识。在范畴化中四种物性特征都起作用,在逻辑转喻中功用和施成信息更重要,在概念转喻中构成的作用更突出,在隐喻中形状的作用更突显。物性结构中百科知识的语言关联还有待于进一步发现与发掘,否则物性角色的数量和明细就无法确定。

百科知识与语言知识还是有界限的,只有符合语言关联原则才能成为语言知识,而语言关联有强弱之分,有系统性与非系统性之分。也许所有的百科知识都会在语言中有所表现,但只有发现系统的语言关联,才能认定为语言知识,才可以纳入词义。比如酒的消毒功能可以帮助我们理解“他把酒洒到伤口上”这样的句子,但并不能由此认定“消毒”就是“酒”词义的一部分。如果不坚持系统性原则,那么必然会得出一切百科知识都属于词义的结论,与认知语言学“百科全书式”的语义观就没什么区别了。当然,怎么界定系统性是个需要探讨的新问题。

一方面,需要确定物性结构中百科知识的语言关联,剔除没有系统语言关联的百科知识。另一方面,物性结构中的知识是否涵盖了具有系统语言关联的所有百科知识也未可知。物性结构特别适合描写指人造物的名词,尤其是处于基本层级的,对指自然物的名词和抽象名词不太适用。例如,“很女人”的理解就无法从“女人”的物性结构中获得。名词语义描写要纳入哪些百科知识及其语言关联性还需要进一步探索。

The Cognitive Orientation of Generative Lexicon: From the Perspective of Linguistic Relevance

Song Zuoyan

(*School of Chinese Language and Literature*, *Beijing Normal University*, *Beijing* 100875, *China*)

Abstract: The research focus in Generative Lexicon Theory (GLT) is on the computational and cognitive modeling of natural language meaning. Both its lexical representation and semantic generative mechanisms are integrated into the cognitive perspective. By contrast, generative lexicon (GL) pays more attention to the linguistic relevance of cognition. The description of word meaning in GL encodes some important encyclopedic knowledge associated with a lexical item through qualia structure, which is the interface of encyclopedic knowledge and linguistic knowledge and enriches the lexical representation of nouns. Such knowledge plays a significant role in semantic composition. Qualia structure is widely applied in the domains of cognitive linguistic study, such as categorization, metaphor, metonymy and conceptual blending, which further reveals the linguistic relevance to encyclopedic knowledge, and shows the similarity between GLT and cognitive linguistics. The linguistic relevance to encyclopedic knowledge in various languages is both universal and relative, which reflects the similarity and dissimilarity between human language and cognition.

Key words: Generative Lexicon Theory, cognitive linguistics, qualia structure, encyclopedic knowledge, linguistic relevance

(责任编辑 管 琴)

① Moravcsik, J. M., *Aitia* as generative factor in Aristotle's philosophy, *Dialogue*, 1975, 14:622-638.